Variation in Glottalization at Prosodic Boundaries in Clear and Plain Lab Speech

Undergraduate Research Thesis

Presented in Partial Fulfillment of the Requirements for graduation "with Honors Research Distinction in Linguistics" in the undergraduate colleges of The Ohio State University

by
Erin Luthern

Ohio State University
April 21, 2015

Project Advisor: Dr. Cynthia Clopper, Department of Linguistics

ABSTRACT

Previous research on glottalization shows that this voice quality occurs more frequently at prosodic boundaries than in the middle of prosodic phrases. This study investigates ten speakers' use of glottalization at prosodic boundaries in five passages read in both clear and plain lab speech. I analyzed each syllable in every passage for its voice quality (glottalized or modal) and for its prosodic boundary strength using the ToBI system. I found that glottalization is used regularly in phrase-final syllables, and that speakers use glottalization marginally more when preceded by a prosodic boundary than when phrase-medial. I found no evidence of a significant effect of speaking style on glottalization use.

ACKNOWLEDGEMENTS

This project would have been impossible without the contributions of some wonderful and obliging people. My thanks go to Rachel Steindel Burdin, for training me in the ToBI system; Christine Prechtel and Shannon Melvin, for their invaluable input and encouragement; and to Cynthia Clopper, for her incomparable guidance and grace. This work was partially supported by an Undergraduate Research Scholarship and the National Science Foundation (BCS-1056409).

1. INTRODUCTION

Glottalization is a voice quality defined by its distinctive acoustic and auditory expression. This quality is reflected in acoustic signals that include irregular periodicity and low fundamental frequency (Redi and Shattuck-Hufnagel, 2010), which can be seen by comparing the aperiodic, glottalized vowel in Figure 1 to the same vowel that is modally voiced in Figure 2. Figure 1 shows clear, individual glottal pulses that are punctuated by breaks throughout the vowel, and each pulse does not look uniform. This voice quality can be contrasted with the modal voicing in Figure 2, which shows closely-spaced pulses that appear generally similar in their size and shape. Glottalization also has a unique auditory impression that is often distinguishable from other voice qualities without visual inspection of the acoustic signal (Redi and Shattuck-Hufnagel, 2010). The auditory impression of glottalization could be described as similar to a creaking door or croaking frog, and has found its way into popular culture to be emblematic of young women in particular. Glottalization defined thus is also known as creaky voice or vocal fry (Henton and Bladon, 1988).

Glottalization as a broad category of voice quality is phonologically contrastive in some languages, such as Chong and Mazatec (Blankenship, 1997), although it is not similarly contrastive in English (Shockey, 2014). Early interpretations of the function of glottalization in English were based on an informal claim that glottalization could signal speaker boredom, and therefore contributed no significant meaning to our understanding of speech (Laver, 1980). However, this interpretation of glottalization was not based on the analysis of actual speech data (Henton and Bladon, 1988).

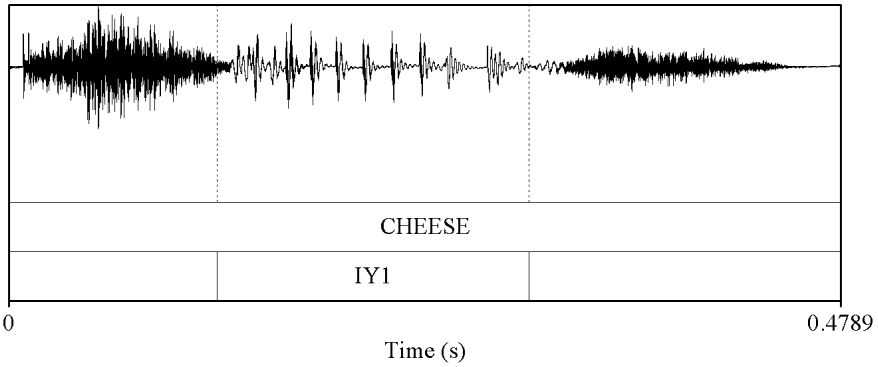

Figure 1: Example of a glottalized syllable “cheese”.

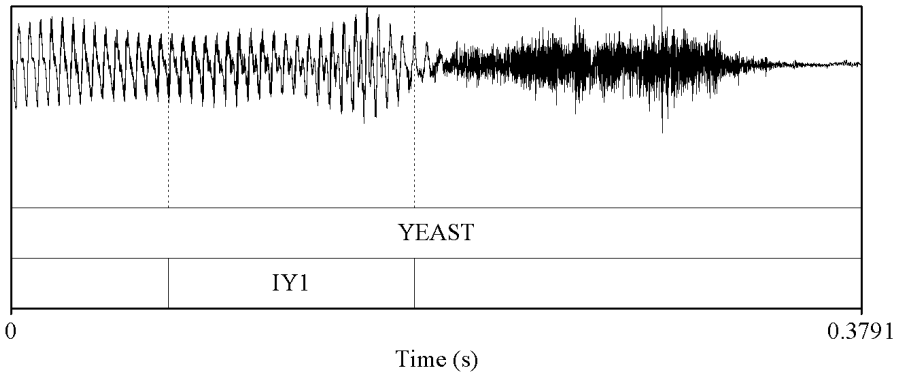

Figure 2: Example of a modal syllable “yeast,” produced by the same speaker as in Figure 1.

More recent research suggests that glottalization is used in English to signal various prominent aspects of speech, including phrase edges and stressed vowels in word-initial position (Garellek, 2014; Kreiman and Sidtis, 2011). Glottalization is associated with phonological processes such as word-final stops becoming glottalized, and the glottalization of vowel-initial words. The process of vowel-initial glottalization is also more likely to occur when the syllable occurs in a

phrase-initial position (Pierrehumbert and Talkin, 1992). Beyond phonology, glottalization is also linked to prosodic patterns, as glottalized syllables are more likely to be observed at phrase boundaries than in phrase-medial position (Dilley et al., 1994; Henton and Bladon, 1988; Redi and Shattuck-Hufnagel, 2001).

Phonetic prominence is also affected by speaking style. For example, clear lab speech, directed towards an imagined hearing-impaired or non-native speaker, has several defining qualities such as an exaggerated pitch range, decreased speaking rate, and exaggerated prominence of pitch accents on accented words in a phrase (Smiljanic and Bradlow, 2009). These enhancements of prominence in clear speech are interpreted as the talker's attempt to assist the listener in correctly parsing the utterance. Given that glottalization is used to mark prominent syllables and prosodic boundaries, I expect that glottalization at prosodic boundaries may be produced more frequently in clear speech than in plain lab speech. The current study tests this prediction through an examination of the frequency of glottalization at prosodic boundaries in clear and plain lab speech.

2. METHODS

2.1. Corpus

The data for this study were collected from an existing corpus of data, containing 30 passages read by 10 male and 20 female college students (Burdin et al., 2010). The subset of data I used was collected from 10 female speakers aged 18-25 years old. All ten speakers are native speakers of the Midland dialect of American English, which is spoken in the southern half of the

American Midwest, stretching from Ohio through Iowa, Missouri, Kansas, and Nebraska, bordered by the Ohio River to the south, and to its northern border including southern and central Ohio, Indiana, and Illinois (Ash, 2006).

Each speaker was prompted to read a series of passages, which were read at a self-paced rate, with each passage displayed on the computer screen one at a time. The five passages selected for this study are provided in the Appendix. All passages were first read in a plain lab speaking style, as though speaking to a friend; speakers were then prompted to read the same passages in a clear lab speaking style, as though speaking to a hearing-impaired or non-native listener. The data therefore comprise a total of 100 passages, with each of the ten speakers reading the five selected passages in both clear and plain speaking styles. Speakers were recorded in a sound-attenuated booth with high-quality digital recording equipment. The read passages were 35-90 s in duration.

2.2. Voice quality coding

Every syllable in the recordings was annotated for voice quality using a series of automated and manual processes. The recordings were first segmented with the Penn Phonetics Lab Forced Aligner (Yuan and Liberman, 2008), and the segmentations were manually checked for accuracy. Based on previous manual coding of voice quality in these data, I assumed an f0 cutoff of 150 Hz, with all f0 values above 150Hz being identified as modal, and those below 150Hz being considered glottalized. Because glottalization can be found at any point within a vowel, I sampled the f0 from each vowel in the passage at 10 equally spaced timepoints throughout the

vowel's duration using Praat (Boersma, 2001). Undefined f0 values were manually corrected; f0 values that were above 350 Hz, and therefore outside the typical range for an adult female speaker, were examined and manually corrected as necessary. Following hand-correction, if six or more of the 10 samples within a given vowel were tagged as modal (with an f0 greater than 150 Hz), the vowel was considered modal. If fewer than six of the 10 samples within a given vowel had an f0 greater than 150 Hz, the vowel was considered glottalized. This automatic method with hand-correction was tested against previous hand-tagged vowels, and an agreement rate of 91% between the automatic and hand-tagged vowels was observed, suggesting that this automated process is highly reliable for identifying glottalization. The total data set included 17,243 tokens, of which 566 were excluded due to text misalignment in the forced-alignment phase or vowel deletion, leaving a total of 16,677 tokens to be analyzed.

2.3. Prosodic annotation

For the analysis of the prosodic boundaries, we used the Tones and Break Indices (ToBI) system (Beckman and Ayers, 1997), which subjectively categorizes the perceived strength of the boundary separating the target word or syllable from those following it. The ToBI scale for break indices creates a 5-point scale from 0 to 4, with the highest values of 3 and 4 being associated with the prosodic constituents of intermediate phrases and intonational phrases, respectively. In the current study, a syllable with a break index of 3 or 4 immediately preceding it was classified as having a preceding prosodic boundary. Similarly, a syllable with a break index of 3 or 4

immediately following it was classified as having a following prosodic boundary. Word-medial breaks between syllables were classified as a 0 break.

3. RESULTS

For each speaker, the proportion of glottalized syllables relative to the total number of syllables was calculated separately for each prosodic position (the presence of a preceding boundary, following boundary, both a preceding and following boundary, or neither) and speaking style (plain vs. clear lab speech). Tokens with both a preceding and following boundary were excluded from the following analysis, as there were only 10 of them. Table 1 shows the mean proportion of glottalized syllables with a preceding boundary, following boundary, or no boundary in both clear and plain speaking styles. A repeated-measures ANOVA with phrase position and speaking style as within-subject factors found that boundary presence was a significant factor in glottalization use ($F(2,18)=9.35$, $p=.002$). This ANOVA also found that speaking style had no significant effect on glottalization use ($F(2,18)=2.71$, $p=.134$), and the interaction of speaking style and boundary presence was also not significant ($F(2,18)=1.64$, $p=.221$). A series of post-hoc paired t-tests revealed the effect of each particular boundary type. A paired t-test comparing preceding boundary to no boundary, collapsed across speaking style showed preceding boundaries led to marginally more glottalization than no boundaries ($t(9)=1.99$, $p=.077$). This result shows that syllables with a preceding boundary are more likely to be glottalized than those that are phrase-medial. A paired t-test also revealed that syllables are more likely glottalized

when followed by a prosodic boundary than when preceded by a boundary ($t(9)=-2.29$, $p=.048$) or phrase medial ($t(9)=4.37$, $p=.002$).

Given the marginal effect of a preceding boundary on glottalization use, I wanted to explore how much of that glottalization came from vowel-initial versus consonant-initial words, given the proposed interaction between vowel-initial words and phrase-initial contexts (Pierrehumbert and Talkin, 1992). A further look at glottalization with preceding boundaries in Table 3 shows the proportions of glottalization in syllables in phrase-initial position when the syllable is consonant-initial or vowel-initial in both speaking styles. A repeated-measures ANOVA revealed that vowel-initial words are more likely to be glottalized when preceded by a prosodic boundary than consonant-initial words ($F(1,9)=20.03$, $p=.002$). The effect of speaking style was not significant ($F(1,9)=2.93$, $p=.12$), nor was the interaction of speaking style and the initial segment ($F(1,9)=2.19$, $p=.17$).

	Clear speech	Plain speech
Following boundary	.48 (.29)	.51 (.25)
Preceding boundary	.22 (.17)	.34 (.24)
No boundary	.16 (.12)	.18 (.11)

Table 1: Proportions of glottalized syllables with a following prosodic boundary, preceding boundary, or no boundary in clear and plain speaking styles across speakers. Standard deviations are shown in parentheses.

	Clear speech	Plain speech
Vowel-initial, phrase-initial	.32 (.20)	.49 (.26)
Consonant-initial, phrase-initial	.17 (.17)	.28 (.27)

Table 2: Proportions of glottalized syllables preceded by a prosodic boundary when the syllable is consonant-initial and vowel-initial across speakers. Standard deviations are shown in parentheses.

One striking aspect of the data that is masked by the statistical analysis is the considerable inter-speaker variation of glottalization between speaking styles with respect to following boundaries, as shown in Table 3. Whereas some speakers showed a large difference in the proportion of glottalization relative to a following intonational phrase boundary in both speaking styles, other speakers did not. For example, Speaker 72 consistently produced glottalization before prosodic boundaries, regardless of speaking style, and is an example of the overall trend towards the importance of following prosodic boundaries in the data. By contrast, Speaker 153, who showed a similar overall degree of glottalization to Speaker 72, exhibited a larger effect of speaking style than prosodic position, using glottalization more frequently in plain speech than clear speech, and is therefore less representative of the sample as a whole.

Although my statistical analysis did not uncover a significant effect of speaking style on glottalization, some speakers showed large stylistic differences that were not captured by the statistical analysis. For example, Speaker 89 glottalized nearly all (90%) of syllables with a following boundary in the clear speaking style, but only 66% of syllables with a following boundary in the plain speaking style. This pattern is consistent with my prediction that glottalization in phrase-final position would be enhanced in clear speech relative to plain speech. However, Speaker 153 showed an increased use of glottalization in plain speech relative to clear

speech, regardless of the presence of a following boundary, contrary to my prediction. Thus, the lack of a speaking style effect in my analysis may reflect individual differences in the use of glottalization across styles.

Speaker	Clear speech		Plain speech		Total
	No F	F	No F	F	
70	0.06	0.68	0.12	0.78	0.15
72	0.24	0.77	0.22	0.67	0.28
73	0.08	0.65	0.07	0.57	0.12
76	0.05	0.41	0.06	0.30	0.09
81	0.14	0.56	0.19	0.56	0.21
86	0.34	0.73	0.31	0.80	0.38
89	0.36	0.90	0.41	0.66	0.43
136	0.16	0.13	0.16	0.11	0.15
150	0.14	0.10	0.14	0.18	0.14
153	0.19	0.14	0.32	0.35	0.25

Table 3: Proportion of individual speakers' use of glottalization with and without a following prosodic boundary in both clear and plain speech, and total proportion of glottalization.

4. DISCUSSION

In this research, I found greater proportions of glottalization at phrase-final boundaries than in non-phrase-final syllables as well as marginally more glottalization at phrase-initial boundaries relative to phrase-medial syllables. The results support the role of glottalization as a marker of prosodic boundaries, even when there is no stop to be glottalized at these boundaries (Redi and Shattuck-Hufnagel, 2001). Previous research suggested that glottalization occurs for various

reasons, whether prosodic or phonological, and that there is some interaction between these functions (Dilley et al., 1994; Henton and Bladon, 1988; Redi and Shattuck-Hufnagel, 2001). My results show that glottalization is used both at phrase-initial and phrase-final positions, which suggests that glottalization is closely linked to suprasegmental prosodic structures, such as phrasing.

I observed marginally significant proportions of glottalization at phrase-initial boundaries than in non-initial syllables, which falls in line with previous research (Dilley et al., 1996). I considered that the proportion of glottalization in phrase-initial position might be strongly affected by the frequency of vowel-initial words at phrase onsets. To explore this possibility, I analyzed phrase-initial syllables, separating them into vowel-initial and consonant-initial categories. The results showed that vowel-initial syllables in phrase-initial position are more likely to be glottalized than their consonant-initial counterparts. These data are in accordance with previous work, which posits that vowel-initial glottalization may be the result of simple physiological difficulty in producing a vowel after a pause in speech (Dilley et al., 1996).

The lack of an effect of speaking style on glottalization suggests that glottalization as a phrase-final marker is used similarly across speaking styles. Thus, although glottalization is used for marking the edges of phrases, this prosodic marking is not exaggerated in clear speech, unlike other dimensions of speech such as pitch range, duration, and vowel quality. An explanation for the variation in glottalization use regarding speaking styles may be found in sociolinguistic variables not accounted for in these data.

The marked variation between speakers' use of glottalization within and across speaking styles suggests that other factors may drive the use of glottalization within this group of speakers. Although an overall effect of speaking style was not observed, some speakers produced the predicted pattern of more glottalization in clear speech than in plain speech, especially before prosodic boundaries (e.g., Speakers 72, 76, and 89). I attempted to control speaker homogeneity by using participants who were all young (18-25 years old), female speakers of the Midland dialect of American English. Because of this control over some basic social factors, I expected to see more uniformity in the overall rates of glottalization. However, my results showed no significant factor to explain the varied use of glottalization between speaking style for individual speakers. For example, Speakers 86 and 89 have similar overall glottalization proportions, but exhibit opposite effects of speaking style. Whereas Speaker 86 shows an increased use of glottalization at phrase-final syllables in plain speech relative to clear speech, Speaker 89 uses much more glottalization at phrase-final syllables in clear speech than in plain speech. The variation in these data suggests that further studies of glottalization and speaking style should explore other potential factors that influence its use.

Further, these data were analyzed from read speech, which may differ from spontaneous speech in terms of glottalization and its prosodic functions. Using read speech allowed me to directly compare identical passages from the same speakers, which allowed for more direct comparisons and provides stronger evidence for individual differences. However, exploration of variation in glottalization across prosodic positions and speaking styles in spontaneous speech

may uncover additional effects that were not observed with the read speech materials in this study.

5. REFERENCES

Beckman, M. E., Ayers, G. 1997. Guidelines for ToBI labelling. The OSU Research Foundation, 3.

Blankenship, B. 1997. The time course of breathiness and laryngealization in vowels. M.A. Thesis, UCLA.

Boersma, P. 2001. Praat, a system for doing phonetics by computer. Glot International 5, 341-345.

Dilley, L., Shattuck-Hufnagel, S., Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. Journal of Phonetics 24, 423-444.

Garellek, M. 2014. Voice quality strengthening and glottalization. Journal of Phonetics 45, 106-113.

Henton, C., Bladon, A. Creak as a sociophonetic marker. In: Hyman, L.M. (ed.), Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin. London: Routledge, 3-29.

Kreiman, J., Sidtis, D. 2011. Foundations of Voice Studies.

Ash, S. 2006. What is dialect? Revisiting the Midland. In: Murray, T.E., Beth, E.L. (eds.), The North American Midland as a Dialect Area. Philadelphia: John Benjamins, 33-56.

Laver, J. 1980. The Phonetic Description of Voice Quality. Cambridge.

Redi, L., Shattuck-Hufnagel, S. 2001 Variation in the realization of glottalization in normal speakers. Journal of Phonetics. 29, 407-429.

Shockey, L. 2014. A personal note on the larynx as articulator in English. In: Waniek-Klimczak, E., Pawlak, M. (eds.), Teaching and Researching the Pronunciation of English. Heidelberg: Springer, 219-222.

Smiljanic, R., Bradlow, A.R. 2009.Speaking and hearing clearly: Talker and listener factors in speaking style changes. Linguistics and Language Compass 3, 236-264.

Yuan, J., Liberman, M. 2008. Speaker identification on the SCOTUS corpus, Proceedings of Meetings on Acoustics 2008. 5687-5690.

APPENDIX

Passage “britain”:

In medieval Britain, food was quite different from how it is now. The majority of people mostly ate bread and cheese. A common drink was mead, which is made from fermented honey. Different wheat and yeast were used in the bread, which made the taste much more inconsistent than we are used to today. The mead - although sweet - would commonly have flies and debris in it, as a side effect of the unsanitary brewing conditions, and the yeast used for brewing was of variable quality. In contrast with the drab cuisine of the peasants, however, food at a royal feast was luxurious. Only the finest wheat would be used in the preparation of sweet pastries, which were often served with imported French cheese. For the main course, meats such as swan or even dolphin were not unheard of. Even today, the right to serve swan meat at a feast is held exclusively by the British royal family.

Passage “chess”:

Nick had made a bet with his sister that he was the best chess player in their local club. They agreed that if he could get to the top of the rankings within two weeks, that would be sufficient proof that Nick was indeed the best chess player. If he failed, though, he owed his sister fifty dollars. In the first week, Nick managed to beat all of his opponents easily, and was very confident that he would succeed. However, the next week he wasn’t able to beat the club’s top player, and was only second place in the rankings. With no definitive proof of his skill, he lost the bet and was fifty dollars poorer.

Passage “flax”:

Clothes made from flax, better known as linen, are durable and with proper care and attention will last a long time. This soft yet strong fabric can be used to make all sorts of clothes and other items. However, care must be taken should a thread come loose. Because flax linen is a soft material, any long thread that comes loose should be immediately clipped off, lest it snag while in the wash. Before putting it in the laundry basket, I recommend making sure the surface is smooth, with nothing to snag on in the wash. The garment is ultimately only as strong as it is smooth.

Passage “mob”:

At the sight of the wounded mob of people desperate to make it through the doors of the makeshift hospital, Amelia felt her chest constrict and a wave of grief and despair wash over her. She had been volunteering there as a nurse for weeks now, but the hospital was chronically understaffed. At the start of the conflict, she would only treat a few patients each day: a man with a bad cough from smoke inhalation, another with shrapnel in his chest, a woman who got hit by stray bullets. Now as the mob slowly moved forward, Amelia straightened her surgical mask, grabbed some gauze, and prepared to act. Even through her mask she could smell the odor of burned skin, and there seemed to be a contagious cough that passed around the assembled crowd. She felt paralyzed with grief at the sight of her people suffering like this, but she knew she had to act. With little more than gauze at hand, she began to treat her patients, her brothers and sisters,

and her friends.

Passage “nature”:

Sandra never really appreciated nature, but her first walk in the countryside really changed her view of the outdoors. A frog startled her by hopping straight across the path in front of her. She found the nest of some animal – probably a shrew – and saw fox footprints too. After climbing to the summit of a hill, there was a spectacular view of the area, especially the diversity of trees. There was even an eagle’s nest in one of them. Now Sandra goes for a walk every Sunday afternoon. The frog by the path no longer startles her. She caught a glimpse of a fox recently, although she’s still never seen a shrew.